EUI WORKING PAPERS IN ECONOMICS

EUI Working Paper ECO No. 95/8

The Fully Modified OLS Estimator as a System Estimator: A Monte-Carlo Analysis

KRISTINA KOSTIAL

European University Institute, Florence

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

ECONOMICS DEPARTMENT

EUI Working Paper **ECO** No. 95/8

The Fully Modified OLS Estimator as a System Estimator: A Monte-Carlo Analysis

KRISTINA KOSTIAL

BADIA FIESOLANA, SAN DOMENICO (FI)

Printed in Italy in March 1995
European University Institute
Badia Fiesolana
I – 50016 San Domenico (FI)
Italy

The Fully Modified OLS Estimator as a System Estimator: A Monte–Carlo Analysis*

Kristina Kostial
European University Institute
e–mail: kostial@datacomm.iue.it

January 1995.

Abstract

This paper provides an extensive Monte–Carlo study on the FM–OLS estimator. The first part analyzes the FM–OLS estimator under alternatives for constructing the long–run variance–covariance matrix. Taking into account six different measures, the prewhitened FM–OLS estimator employing the Bartlett kernel with an automatic choice of the bandwidth parameter proves to be the best. Section two focuses on the sensitivity of the estimator to non–Gaussian error processes; especially autoregressive error processes with roots close to unity can constitute a serious problem.

The following sections are of particular relevance for applied work. Section three illustrates that the higher the rank of the cointegrating space, the more precise will be the estimates with the FM–OLS estimator. The next section highlights the possibility of exploiting permutations of the data set to determine the rank of the cointegrating space. Finally, the results of the Monte–Carlo study are applied to estimating a data set with the FM–OLS estimator.

KEYWORDS: Cointegration, Monte–Carlo Analysis, Fully Modified Estimation.

*I would like to thank Niels Haldrup and Grayham E. Mizon for very helpful discussions on the issues developed in the paper. Financial support from the European Investment Bank is gratefully acknowledged. Of course, the usual caveat applies.

Contents

1 Introduction

A vast literature offers numerous estimators for cointegrating relationships. Most of these estimators are single equation estimators. Amongst those the best known estimators are: the OLS as proposed in Engle and Granger (1987), the Three–Step Estimator by Engle and Yoo (1991), and Augmented Least Squares Approaches as suggested by Stock (1987), Bårdsen (1988) or Saikonnen (1991).[1] In order to analyze a system, there exist two methods on principle. One method is based on finding eigenvalues and eigenvectors of appropriately defined matrices. This technique is exploited by the widely used Johansen estimator [Johansen (1988), (1991)], or the Modified Box–Tsiao estimator as proposed by Bewley, Orden and Fisher (1991). The other method for analyzing a system is the so–called Fully Modified OLS estimator, a semi–parametric approach, which is suggested by Phillips (1991a), and Phillips and Hansen (1991).

The logical consequence of a variety of alternative estimators is a multitude of Monte–Carlo studies in order to analyze the the weaknesses and strengths of the different estimators. Most of the Monte–Carlo studies compare the behavior of several estimators in different situations, see e.g. Cappuccio and Lubian (1992), Hargreaves (1994), and Inder (1991, 1993). Others concentrate on one estimator, e.g. Cappuccio and Lubian (1994), who focus on the Fully Modified OLS estimator.

Concerning the use of the Fully Modified OLS estimator, these studies have one common characteristic. Although the theoretical set–up of the Fully Modified OLS estimator allows for a system analysis, the estimator is only used as a single–equation estimator. This paper analyzes the qualities of the Fully Modified OLS estimator as a system estimator. A four dimensional system of cointegrated $I(1)$ variables is considered to find answers to the following questions:

- Does the use of different kernels for the non–parametric estimation

[1]This list is incomplete and could be augmented by methods using instrumental variables [Phillips and Hansen (1990)], canonical cointegration regression [Park (1992)], or spectral regression [Phillips (1991b)].

of the long–run variance–covariance matrix matter?

- How does prewhitening influence the results?

- How important is the choice of the bandwidth parameter?

- How sensitive is the Fully Modified OLS estimator to non–Gaussian error processes?

- Is the performance of the FM–OLS estimator dependent on the relation of the rank of the cointegrating space to the dimension of the system?

- In the case of more than one cointegrating relationship: Can permutations of the data set help to determine the rank of the cointegrating space?

The remainder of the paper is organized as follows: After an introduction of the Fully Modified OLS estimator in section 2, section 3 contains the design of the Monte–Carlo experiment. Section 4 is dedicated to the search for the "optimal" Fully Modified OLS estimator. Herein included is the discussion of the appropriate kernel for the long–run variance–covariance matrix, the benefit from prewhitening versus non–prewhitening, and the influence of the bandwidth parameter.

The further analyses in sections 5 to 8 are conducted with the estimator, which proves to be "optimal" under the aspects of the preceding section. In section 5 the the Fully Modified OLS estimator is exposed to different DGPs in order to analyze its sensitivity to non–Gaussian error processes and different kinds of exogeneity. Section 6 focuses on the role of the rank of the cointegrating space for the quality of the estimator. The possibility of determining the rank of the cointegrating space by estimating permutations of the variables is reported in section 7. In section 8 the usefulness of the Fully Modified OLS estimator as a system estimator is illustrated by an application to a data set as estimated in Clements and Mizon (1991) with the Johansen procedure. Section 9 concludes.

2 The Fully Modified OLS Estimator

In order to summarize the estimation method of Phillips (1991a) and Phillips and Hansen (1990), it is convenient to adopt the notation of Phillips (1991a). He suggests a triangular system representation for a N dimensional system of $I(1)$ variables with r cointegrating relationships:

$$\Delta X_t = -EAX_{t-1} + v_t, \tag{1}$$

where

$$E := \begin{pmatrix} I_r \\ 0 \end{pmatrix} \text{ and } A := \begin{pmatrix} I_r & -B \end{pmatrix}. \tag{2}$$

The triangular system representation for the case, in which the error process v_t is stationary, can be estimated parametrically [Phillips (1991a)][2] or by a the so–called Fully Modified OLS estimator, a semiparametric correction proposed by Phillips and Hansen (1990).

Phillips and Hansen decompose the long-run variance–covariance matrix Γ of the error process v_t into

$$\begin{aligned} \Gamma &= \Omega + \Lambda + \Lambda' &(3)\\ &:= E(v_1 v_1') + \sum_{k=2}^{\infty} E(v_1 v_k') + \sum_{k=2}^{\infty} E(v_k v_1'). &(4) \end{aligned}$$

They then use consistent estimates of Ω, Λ and Γ to construct the Fully Modified OLS estimator:

$$\begin{aligned} \hat{B} = & \left[\sum_1^T X_{1t}^+ X_{2t}' - T\left[I_r, -\hat{\Gamma}_{12}\hat{\Gamma}_{22}^{-1}\right]\left[\hat{\Omega}_{21} + \hat{\Lambda}_{21}, \hat{\Omega}_{22} + \hat{\Lambda}_{22}\right]'\right] \\ & \left[\sum_1^T X_{2t} X_{2t}'\right]^{-1}. \end{aligned} \tag{5}$$

In equation (5), the superscript $\hat{}$ denotes the consistent estimates, $X_{1t}^+ := X_{1t} - \hat{\Gamma}_{12}\hat{\Gamma}_{22}' \Delta X_{2t}$, and Γ, Λ and the Ω matrices are partioned conformably

[2]Following the approach of Dunsmuir and Hannan (1976) and Dunsmuir (1979), a good parametric approximation for the Gaussian Likelihood is given by the Whittle Likelihood.

with X_t. Under certain regularity assumptions on the error process v_t this semiparametric procedure is asymptotically equivalent to full maximum likelihood.[3]

The crucial aspect of the Fully Modified OLS procedure is the consistent estimation of the elements of the long–run variance–covariance matrix. The usually employed method, which guarantees consistency, is the kernel estimation as offered by Andrews (1991). This estimation method requires the choice of a kernel and a bandwidth parameter. Any kernel that yields positive semidefinite estimates can be used; included in this set are the truncated, the Bartlett, the Parzen, the Tukey–Hanning and the quadratic spectral kernel.

The choice of the bandwidth parameter is more complicated. Andrews regards a bandwidth parameter as being optimal if it minimizes the asymptotic truncated mean squared error. He can show that the rate of convergence for the optimal bandwidth parameter is $T^{1/3}$ for the Bartlett kernel and $T^{1/5}$ for the Parzen, Tukey–Hanning and quadratic spectral kernel. Furthermore, he recommends a plug–in bandwidth estimator, which he derives by an approximation of parametric AR(1) models for each element of the error process.[4]

Since in most of the applications the cointegrating residuals v_t display a significant degree of correlation, Hansen (1992) suggests to prewhiten the residuals by a VAR(1):

$$\hat{v_t} = \Phi \hat{v_{t-1}} + \hat{\epsilon_t}. \tag{6}$$

[3]Phillips and Hansen impose several assumption on the innovation error process v_t. It has to be strictly stationary and ergodic with zero mean, finite covariance matrix Ω, and continuous spectral density matrix $f_{vv}(\lambda)$ with $\Gamma = 2\pi f_{vv}(0)$. Furthermore, the partial sum process constructed from the error process is supposed to satisfy the multivariate invariance principle

$$T^{-1/2} \sum_{1}^{[Tr]} v_j \qquad \Longrightarrow \qquad B(r) \equiv BM(\Gamma), \qquad\qquad 0 < r \leq 1.$$

[4]Section 4.1 and 4.3 provide more details on the kernels and the choice of the optimal bandwidth parameter.

He then applies the kernel estimation to the whitened residuals $\hat{\epsilon}$. Consistent estimates[5] of the covariance parameters of interest can be obtained by recoloring:

$$\hat{\Gamma} = (I - \hat{\Phi})^{-1} \hat{\Gamma}_{\epsilon} (I - \hat{\Phi}')^{-1} \tag{7}$$

$$\hat{\Omega} + \hat{\Lambda} = (I - \hat{\Phi}')^{-1} (\hat{\Omega}_{\epsilon} + \hat{\Lambda}_{\epsilon}) - (I - \hat{\Phi})^{-1} \hat{\Phi} T^{-1} \sum_{i=1}^{T} \hat{v}_t \hat{v}_t{}' . \tag{8}$$

3 The Design of the Monte–Carlo Experiment

In order to compare the results of the system analysis to those obtained from the single–equation case, the design of the Monte–Carlo experiment by Hargreaves (1994) was adopted. Hargreaves sets up a four dimensional model, which is given by:[6]

$$B z_t = u_t \qquad \text{with} \qquad u_t = A^1 u_{t-1} + A^2 u_{t-2} + \epsilon_t + \Theta \epsilon_{t-1}. \tag{9}$$

The matrices B, A^1, A^2 and Θ are all 4×4, and B of full rank. The error process ϵ_t is $iidN(0, \Omega)$. The model defined above allows for a huge variety of ARIMA error processes and can have up to three cointegrating vectors. Rewriting the model in (9) in ECM form yields

$$\Delta z_t = -B^{-1}(I_4 - A^1) B \Delta z_{t-1} - B^{-1}(I_4 - A^1 - A^2) B z_{t-2} + \epsilon_t + \Theta \epsilon_{t-1} \tag{10}$$

such that the long–run matrix is given by

$$\Pi = -B^{-1}(I_4 - A^1 - A^2) B. \tag{11}$$

Various standard measures are used to analyze the simulation results. Amongst those are the mean bias, the standard deviation, skewness, and kurtosis. The measures for skewness and kurtosis are drawn

[5] The consistency is proved in Andrews and Monahan (1992).

[6] See Hargreaves (1994) for an embedding of the DGPs used by Hansen and Phillips (1990) and Phillips and Loretan (1991).

from Bewley, Orden and Fisher (1991). Defining q_i as the ith quartile, skewness s is given by

$$s := \frac{q_{97.5} - q_{50.0}}{q_{50.0} - q_{2.5}} - 1, \tag{12}$$

and kurtosis k by

$$k := \frac{q_{97.5} - q_{2.5}}{q_{99.5} - q_{0.5}} - \frac{1.96}{2.575}. \tag{13}$$

Both measures equal zero for the standard Normal distribution. In order to investigate the outlier behavior, the minimal and maximal values of the estimators were computed. Furthermore, the quartiles for 2.5, 5.0, 10.0, 50.0, 90.0, 95.0, and 97.5 percent are supplied.[7]

The Monte–Carlo experiments were programmed using the GAUSS-v.3 package (with its random number generator) and 5.000 replications of each experiment were performed. The number of observations was set to 100.

4 The Choice of the "Optimal" Fully Modified OLS Estimator

As described in the section 2, the crucial ingredients for the Fully Modified OLS estimator are the consistent estimates of the elements of the long–run variance–covariance matrix. For this Monte–Carlo study the kernel estimation as proposed by Andrews (1991) is applied. This leaves the choice of the appropriate kernel and the bandwidth parameter. Furthermore, there is a third choice involved in the construction of the long–run variance–covariance matrix. This is whether to apply the procedure of prewhitening to the error process or whether to use the "raw" error process.

This section analyzes the influence of the three alternatives on the quality of the estimates derived with the Fully Modified OLS estimator. The goal is to find a combination of the these alternatives, which

[7] The reported tables and graphs are a condensed version of the available information.

generates a Fully Modified OLS estimate which is reliable under various designs of the error process.

It is an unrealizable task to find a set of models which represents all the possible DGPs in the economy, which an econometrician has to face. Nevertheless, one can simulate different specifications in order to analyze the reaction of the estimators. This Monte–Carlo experiment consists of eighteen different DGPs, which are oriented on the choice of Hargreaves (1994) and originated from model (9).[8] The following were taken as constants:

- $$\begin{pmatrix} B_{13} & B_{14} \\ B_{23} & B_{24} \\ B_{33} & B_{34} \\ B_{43} & B_{44} \end{pmatrix} \equiv \begin{pmatrix} 5/7 & -1/7 \\ 7/4 & 1/4 \\ -2/3 & 1/3 \\ 3 & 1 \end{pmatrix},$$

- $$\begin{pmatrix} A^1_{13} & A^1_{14} \\ A^1_{23} & A^1_{24} \\ A^1_{33} & A^1_{34} \\ A^1_{43} & A^1_{44} \end{pmatrix} \equiv \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1.5 & 1 \\ 0 & 1 \end{pmatrix},$$

- $$A^2 \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -0.5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The specification of A^1 and A^2 makes u_{3t} and u_{4t} to an autoregressive integrated (moving average) process and a random walk (with moving average part), respectively. The setting of the other parameters, B_{ij}, A^1_{ij}, $(i = 1, \ldots, 4, j = 1, 2), \Theta$, and Ω, is displayed in table 1 and is chosen such that every DGP has two cointegrating vectors. Note that model (9) is set up such that weak exogeneity of z_{3t} and z_{4t} implies no Granger causality from the first two variables to z_{3t} and z_{4t}, i.e. weak exogeneity implies strong exogeneity. DGP 12 considers the case of Granger causality without the presence of weak exogeneity.

[8] Note that these DGPs were chosen as being representative from a larger set of experiments.

Although many more different simulations could have been performed, the DGPs generated by the settings in table 1 cover a multiplicity of different processes. Included are changes in the variance–covariance matrix of the error process (DGP 2, 3, 4, 11), and changes in the B matrix (DGP 7 to 12), which should help to clear the role of exogeneity for the results. Furthermore, the influence of autoregressive (DGP 5, 6, 10) and moving average error processes (DGP 13 to 18) is analyzed. It is clear that an evaluation based on eighteen DGPs gives certain weights to certain situations. But a close look at the tables can confirm that the behavior of the estimators is essentially uniform over the different DGPs and that there are no outliers associated with specific types of DGPs.

Since a variation of all three alternatives would lead to an insurmountable amount of tables, the first two sections use the automatic bandwidth parameters as proposed by Andrews (1991). Section 4.1 takes into consideration five different kernels for the prewhitened as well as the "raw" error processes to enter the formulation of the long–run variance–covariance matrix. Since the estimates attained by one of the five kernels are quite badly behaved, section 4.2 concentrates on the characteristics of the estimates using four different kernels under the aspect of prewhitening versus the "raw" estimator. Exploiting the results of the preceding section 2, section 4.3 can focus on the effect of the choice of fixed and automatic bandwidth parameters on the estimates.

4.1 The Choice of the Appropriate Kernel

The idea of nonparametric or kernel estimation is to approximate the long–run variance–covariance matrix Γ as 2π times an estimate of the spectral density matrix at frequency zero:

$$\hat{\Gamma} = \sum_{j=-T}^{T} k(j/S_T)\hat{\Upsilon}(j), \tag{14}$$

where

$$\hat{\Upsilon}(j) := \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \hat{u}_t \hat{u_{t-j}}' & \text{for } j \geq 0, \\ \frac{1}{T}\sum_{t=-j+1}^{T} \hat{u_{t+j}}\hat{u}_t' & \text{for } j < 0. \end{cases} \tag{15}$$

The kernel $k(x)$ indicates how much weight is given to each frequency; the so-called bandwidth parameter S_T determines the frequencies, which are viewed to be necessary to enter the estimate of the spectral density matrix evaluated at zero frequency.

For the choice of the appropriate kernel, Monte–Carlo experiments including the Fully Modified OLS estimator have always been guided by the propositions of Andrews (1991). Andrews focusses on five different kernels in his article: the truncated, the Bartlett, the Parzen, the Tukey–hanning and the quadratic spectral kernel.[9] Although he recommends the quadratic spectral kernel, many simulation studies used other kernels for the estimation of the long–run variance–covariance matrix. See e.g. Hargreaves (1994), who employs the Parzen kernel, or Lubian and Cappuccio (1994), who consider the Bartlett kernel besides the quadratic spectral kernel, or Hansen (1992), who estimates with the Bartlett, the Parzen and the quadratic spectral kernel with nearly identical results.

This part of the Monte–Carlo study is concerned with the influence of the use of different kernels on the behavior of the Fully Modified OLS estimator. Since detailed tables reporting every result for every DGP are equivalent to not seeing the wood for trees, graphs will try to illustrate

[9] The kernels are defined as:

$$\text{Truncated kernel:} \qquad k_{TR}(x) := \begin{cases} 1 \text{ for } |x| \leq 1, \\ 0 \text{ otherwise,} \end{cases}$$

$$\text{Bartlett kernel:} \qquad k_{BA}(x) := \begin{cases} 1 - |x| \text{ for } |x| \leq 1, \\ 0 \qquad \text{otherwise,} \end{cases}$$

$$\text{Parzen kernel:} \qquad k_{PA}(x) := \begin{cases} 1 - 6x^2 + 6|x|^3 & \text{for } 0 \leq |x| \leq 1/2, \\ 2(1-|x|)^3 & \text{for } 1/2 \leq |x| \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{Tukey–Hanning kernel:} \qquad k_{THK}(x) := \begin{cases} (1+\cos(\pi x))/2 & \text{for } |x| \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{Quadratic Spectral kernel:} \qquad k_{QS}(x) := \frac{25}{12\pi^2 x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right).$$

For a detailed discussion see Andrews (1991).

the main consequences.[10] The estimators were evaluated taking into account six criteria, i.e. their bias and standard deviation, the 5 and 95 percent quartiles and their skewness and kurtosis. In order to summarize the behavior of the estimators, the following grading rules were applied separately to each of the four coefficients of the estimated cointegrating relationships and to each of the eighteen DGPs:

- Basically, ranks from one to five are allocated to the estimators employing the different kernels. Number one is related to the best, and number five to the estimator, which performed worst.
- The criteria for determining the ranks were:[11]
  - 3rd digit after the decimal point for the bias,
  - 5th digit after the decimal point for the standard deviation,
  - 2nd digit after the decimal point for the 5 and 95 percent quartiles and skewness and kurtosis.
- If estimators achieve the same figure for one of the criteria, they get the same rank for that criterion. The next best estimator will be allocated the next best number, if the figure for the criterion does not differ significantly (in the sense of differing in the $k-2$th digit if the kth digit is chosen as the criterion). In the case of significantly worse behavior, the next best estimator receives mark five.

Graphs in the appendix display the average taken over each estimated coefficient and each DGP for the six criteria. First consider graph 1 for the "raw" estimators. It shows quite clearly that the quadratic spectral kernel is very badly behaved; in all but the skewness and kurtosis criteria it is the worst choice. The truncated kernel is the next best choice, but it is outperformed by the triple formed by the Bartlett, the Parzen, and the Tukey–Hanning kernel estimators, which act very similarly. Especially, their bias–behavior and the 90 percent band, which is formed by the 5 and 95 percent quartiles, are very favorable.

[10] Detailed tables are available on request from the author.

[11] Always in relation to the true parameter value.

Next turn to graph 2 for the prewhitened estimators. There appears to be a an underlying pattern, which is common to the two graphs. Again, the quadratic spectral kernel is widely surpassed by the other estimators; its only advantage seems to lie in a small skewness and thin tails. The four other estimators form a group with similar behavior. Nevertheless, the results for the truncated and the Bartlett kernel seem to be slightly more auspicious.

As mentioned before, Andrews recommends the use of the quadratic spectral kernel since it is optimal with respect to asymptotic truncated mean squared error within a certain class of kernels, which includes all but the truncated kernel. In the Monte–Carlo experiments, though, the performance of the quadratic spectral kernel does not confirm the optimality of this kernel; on the contrary: the choice of the quadratic spectral kernel seems to be the worse choice amongst the five kernels for estimating with the Fully Modified OLS estimator. This result is not necessarily an inconsistency. The asymptotic truncated mean squared criterion used by Andrews is justifiable in the context of constructing standard errors or variance estimators for simple OLS estimates. If, on the other hand, the estimate of the long–run variance–covariance matrix is used to form more complicated test–statistics or estimates like in the case of the Fully Modified OLS estimator, the suitability of the truncated mean squared error criterion is less clear. This might happen for cases, in which deviations of one part of a test statistic from its limiting behavior can be offset by deviations from another part of the statistic from its limiting behavior in small samples. In such cases, the argument for the optimality of the quadratric spectral kernel breaks down.

4.2 Prewhitening versus the "Raw" Estimator

Most of the recent Monte–Carlo studies and applications of the Fully Modified OLS estimator have adopted the idea of prewhitening the estimated residuals before applying the kernel estimation. This procedure is confirmed by an extensive Monte–Carlo study by Cappuccio and Lubian (1994). They strongly favor the use of VAR prewhitening in single–equation estimation with the Fully Modified OLS estimator. Nevertheless, this section wants to check whether the results derived for the single–equation framework can be transferred to system estimation and whether they depend on the employed kernel.

Since the performance of the estimates derived by using the quadratic spectral kernel was quite disappointing, the analysis is concentrated on the other kernels. Applying the same criteria for determining the ranks of the estimators as introduced in the preceding section, graph 3 displays the number of the estimates over the different DGPs and coefficients, for which the prewhitened estimator received better results than the "raw" estimator.

Taking into account that the maximal achievable number is 72 (which is 18 DGPs times 4 coefficients), there does not seem to exist an unequivocal advantage of the either of the procedures. In fact, there are many cases, in which none of the procedures proves to be more favorable than the other.[12] The next observation, which hits the eye, is that the distribution of the prewhitened estimators is more skewed but has much thinner tails than that of the "raw" estimators.[13] Or, in other words, the procedure of prewhitening seems to have a positive effect on the outlier behavior of the estimator.

Next turn to the graphs for the particular kernels. A look at the graph for the truncated and the Bartlett kernel shows clearly that the prewhitened estimator is superior to the "raw" estimator. For the truncated kernel this result is not unexpected, since the truncated kernel was not in the triple of the "optimal kernels" for the "raw" estimator, but one of the "optimal kernels" for the prewhitened estimator. Only the results for the skewness and kurtosis are less favorable for the prewhitened estimator.

The advantage of the prewhitened estimator against the "raw" estimator is not as distinct for the other kernels. The Tukey–Hanning kernel, though one of the "optimal kernels" for the "raw" estimator but

[12] This cannot be concluded from the graphs, but from further background information.

[13] A result, which is similar to that of Cappuccio and Lubian (1994).

not for the prewhitened estimator, has slightly more favorable results for the prewhitened estimator. The Parzen kernel, finally, is found to be better behaved for the "raw" estimator. The clear superiority of the prewhitened estimator at the 5 percent quartile respectively the clear superiority of the "raw" estimator at the 95 percent quartile is due to the differently skewed biases of the alternative estimators.

Another item meriting attention is the question whether the performance of the prewhitened and "raw" estimator does depend on the type of the underlying error process. (For clarity of presentation, the tables, which could answer these questions, are condensed in the graphs.) Indeed, there seems to exist a weak link between the error process and the performance of the different estimators. Not unexpectedly, the prewhitened estimators produce slightly better results for autoregressive error processes and significantly better results for moving average error processes with roots close to unity.

Summarizing the results above and confirming the extensive Monte–Carlo study for single–equation estimation by Cappuccio and Lubian (1994), it can be concluded that the prewhitened Fully Modified OLS estimator produces on average better results than the "raw" estimator. Nevertheless, the decision for or against prewhitening should be taken after the choice of the kernel. If the truncated of the Bartlett kernel will be applied, the prewhitening of the residuals is appropriate. If, on the other hand, the Parzen kernel is chosen, the alternative seems to be more promising. The remainder of the paper will be conducted with the prewhitened Fully Modified OLS estimator using the Bartlett kernel.

4.3 The Influence of the Bandwidth Parameter

There exist two different ways to choose the bandwidth parameter for the kernel estimation of the long–run variance–covariance matrix. The theoretical framework for both procedures is due to Andrews (1991). One way is simply to use a fixed bandwidth parameter, which is below the rate of convergence for the optimal bandwidth parameter[14] as determined by

[14] Optimal in the sense of minimizing the asymptotic mean squared error.

Andrews. The other way is to follow the recommendation of Andrews for an automatic bandwidth estimator.

In order to derive the automatic bandwidth parameter for the Bartlett kernel, univariate AR(1) models for each component of the error process v_t are considered. Let $(\hat{\rho}_i, \hat{\sigma}_i)$ denote the autoregressive and innovation variance estimates for the ith element of the error process. The automatic bandwidth parameter for the Bartlett kernel $\hat{S}_T$ is then given by

$$\hat{S}_T = 1.1147 \left(\sum_{i=1}^{4} \frac{4\hat{\rho}_i^2 \hat{\sigma}_i^2}{(1-\hat{\rho}_i)^6 (1+\rho_i)^2} \Big/ \sum_{i=1}^{4} \frac{\hat{\sigma}_i^2}{(1-\hat{\rho}_i)^4} T \right)^{1/3}, \qquad (16)$$

where T is the number of observations. Instead of the univariate AR(1) models, the parametric approximation can as well be performed by a VAR(1); an approach, which is adopted by Cappuccio and Lubian (1994). The univariate approximation has the advantage of simplicity and parsimony over the use of a single multivariate model and therefore is given preference in this Monte–Carlo experiment.

Since plug–in methods are characterized by the use of an asymptotic formula for an optimal bandwidth parameter, the goal of this section is to investigate whether the automatic bandwidth is significantly superior to a fixed bandwidth parameter in a small sample size. Two different DGPs will be considered, using the prewhitened Fully Modified OLS estimator with the Bartlett kernel. The chosen DGPs are the "benchmark" DGP 1, and DGP 6, a DGP with an autoregressive error with a root close to one, for which the Fully Modified OLS estimator does not perform very convincingly.[15]

Table 2 displays the bias, standard deviation, 5 and 95 percent quartiles, and skewness and kurtosis for the automatic bandwidth parameter and several fixed bandwidth parameters. The cointegrating vectors result from the estimates in the tables as:

$$\begin{pmatrix} 1 & 0 & c_{11} & c_{12} \\ 0 & 1 & c_{21} & c_{22} \end{pmatrix}.$$

[15] As can be seen in the next section.

The outcomes for DGP 1 do not differ significantly under the choice of automatic or fixed bandwidth; though they seem to be slightly more favorable for a small fixed bandwidth parameter ($S = 3$). Similarly, DGP 6 does not surprise with huge differences between the various bandwidths. Nevertheless, the distribution of the estimate using the automatic bandwidth parameter appears to be less distorted. Summing up it might be said that neither of the procedures turns up to be dominant. It should be born in mind, though, that the automatic bandwidth parameter releases the econometrician from the responsibility of choosing the "right" bandwidth.[16]

5 Non–Gaussian Error–Processes

After having dealt with the problem of finding the Fully Modified OLS estimator, which performs best on average under a variety of DGPs, this section focuses on the sensitivity of the Fully Modified OLS estimator to non–Gaussian error processes and exogeneity. The same DGPs, over which the average was taken to choose the "optimal" Fully Modified OLS estimator, are now considered in more detail. The choice of the DGPs is mainly oriented by Hargreaves's (1994) Monte–Carlo experiments and therefore allows for a comparison to the single–equation estimation with the Fully Modified OLS estimator.[17]

Based on the performance of the estimator for the benchmark DGP 1 (subsection 5.1), the Fully Modified OLS estimator will be evaluated for four different situations. Subsection 5.2 analyzes the influence of the variance–covariance matrix of the error process, subsection 5.3 the effect

[16] Note that this experiment considered two DGPs only and thus cannot be regarded as representative. An argument in favor of the representativeness of the results, though, is that two extreme cases are chosen for the DGPs such that one leads to reliable estimates with the Fully Modified OLS estimator whereas for the other the estimates derived with the Fully Modified OLS estimator are rather disappointing.

[17] The only difference between Hargreaves' DGPs and the DGPs in this Monte–Carlo experiment is that Hargreaves has an integrated autoregressive error process in the second equation. This choice sets the rank of the cointegrating space equal to one.

of changes in the B matrix. DGPs with autoregressive and moving average error processes are scrutinized in subsection 5.4 and 5.5, respectively. Subsection 5.6 concludes.

Guided by the results in section 4, tables 3 to 5 report the relative bias and standard deviation, the 5 and 95 percent quartiles, and skewness and kurtosis for the prewhitened Fully Modified OLS estimator employing the Bartlett kernel with an automatic bandwidth choice.[18]

5.1 The Benchmark DGP (DGP 1)

The tables display quite impressive results for the relative bias and standard deviation. Similarly, the 90 percent band, which is spanned by q05 and q95, is relatively close around the mean. The distribution of the Fully Modified OLS estimator is significantly skewed to the right, especially for the estimates of the second cointegrating relationship. This might be due to the fact that the Monte–Carlo mean of the estimates for all coefficients but coefficient c_{12} lies to the left of the true value. The kurtosis is small and negative for all coefficients; thus, the distribution of the estimator has thinner tails than the standard normal distribution. Note that the significantly worse behavior of the estimate of coefficient

[18] The normed cointegrating vectors of the DGP for the DGPs 1 – 6, 9, 12 to 18 are

$$\begin{pmatrix} 1 & 0 & 1 & -0.035 \\ 0 & 1 & 1 & 0.379 \end{pmatrix},$$

for the DGPs 7 and 8 are

$$\begin{pmatrix} 1 & 0 & 0.714 & -0.143 \\ 0 & 1 & 2.333 & 0.333 \end{pmatrix},$$

for DGP 10 are

$$\begin{pmatrix} 1 & 0 & 1.381 & -0.048 \\ 0 & 1 & 2.333 & 0.333 \end{pmatrix},$$

and for DGP 11 are

$$\begin{pmatrix} 1 & 0 & 0.714 & -0.143 \\ 0 & 1 & 1.381 & 0.524 \end{pmatrix}.$$

c_{12} in relation to the estimates of the other coefficients might be due to the very small value of the coefficient.

5.2 Variance–Covariance Matrix of the Error Process (DGPs 2 to 4, and 11)

The impact of changes in the variance–covariance matrix of the error process on the quality of the Fully Modified OLS estimator is subject of this subsection. Three DGPs are considered, for which the variance–covariance matrix has been changed with respect to the benchmark DGP 1.

25–fold increase of the variance of the error ϵ_{2t} for the second cointegrating relationship (DGP 2) Not unexpectedly, the estimates for the second cointegrating relationship are not as good as the estimates in the benchmark case.[19] Although the relative biases almost remain the same, the relative standard deviation has increased and the 90 percent band (as can be computed with q05 and q95) is significantly widened. The skewness, though, is much smaller, i.e. the distribution is still skewed to the right, but not as much as for the benchmark DGP. The kurtosis is either the same (coefficient c_{21}) or slightly higher, but still negative (coefficient c_{22}). These results confirm the outcome of the single–equation analysis with respect to relative bias and standard deviation, but not for the measures of the distortion of the distribution. For the single–equation estimation, the effect on skewness and kurtosis are significant increases for a higher variance in the error.

Surprisingly, the estimates for the first cointegrating relationship seem also to be affected by the change in the variance–covariance matrix: the relative biases and the standard deviation are smaller and the 90 percent band narrower. The skewness is significantly reduced and the

[19]Note that a normation of the first two equations leads to

$$\begin{pmatrix} 1 & 0 & 1 & -0.035 \\ 0 & 1 & 1 & 0.379 \end{pmatrix} z_t = \begin{pmatrix} 0.724 & 0.276 & 0 & 0 \\ -0.966 & 0.966 & 0 & 0 \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}. \qquad (17)$$

distribution remains being twisted to the right; the kurtosis is almost the same.

Positive Intercorrelation of Errors (DGP 3) For this DGPs the correlation matrix of the errors was set so that $corr(u_{it}u_{jt}) = 0.8^{|i-j|}$. Relative bias and standard deviation of the Fully Modified OLS estimator react negatively to the correlation in the errors. Accordingly, the 90 percent range created by the 5 and 95 percent quartiles is slightly wider. The effect on the skewness is not clear; in fact, DGP 3 is one of the few DGPs, for which the distribution of the estimator for some parameters is skewed to the left. Although the skewness appears to be highly effected by the correlation of the errors, the kurtosis displays minimal changes.

Negative Intercorrelation of Errors (DGP 4) DGP 4 has a contemporaneous negative correlation of the errors such that $corr(u_{it}u_{jt}) = (-0.8)^{|i-j|}$. The effect of the negatively correlated errors on relative bias, standard deviation, and 90 percent band are not as strong as that of the positively correlated errors in DGP 3. Surprisingly, these measures are minimally better for the estimates of the second cointegrating relationship. Nevertheless, a minor loss of quality of the estimates for the first cointegrating relationship has to be noticed. There are changes in skewness and kurtosis, but they do not seem to follow a particular pattern. Hargreaves (1994) performed Monte-Carlo simulations similar to DGP 3 and 4, but for one cointegrating relationship. Analogously, he reports a more significant impact of the positively correlated errors on the estimates.

Mixed Intercorrelation of Errors (DGP 11) DGP 11 is a modification of DGP 9, which is in triangular system form. The errors in DGP 11 are negatively as well as positively correlated with each other. In comparison to DGP 9, bias, variance, 90 percent band, and kurtosis remain almost unchanged; merely the skewness has increased.

5.3 Changes in the B Matrix (DGPs 7 to 12)

For the influence of changes in the B matrix on the estimation with the Fully Modified OLS estimator, six different DGPs will be considered. Main issue is to analyze whether the represention of the DGP in triangular system form with Gaussian errors improves the quality of the Fully Modified OLS estimator. Furthermore, some remarks on the validity of weak and strong exogeneity are added. DGP 1 serves as benchmark.

Perfect Triangular System Representation (DGPs 9 to 11) DGP 9 represents a DGP, which should be ideal for the Fully Modified OLS estimator. Not only that the DGP is in triangular system from, but also the regressors z_{3t} and z_{4t} are strongly exogenous for both cointegrating relationships in DGP 9. Comparing DGP 9 with benchmark DGP 1, it has to be admitted that, though DGP 9 seems to simulate an ideal situation, the Fully Modified OLS estimator does not respond to this improvement; the results are as good as for the benchmark DGP. Note that the estimates for the parameters for the second cointegrating relationship are still significantly skewed to the right.

DGP 11 has already been discussed before, and DGP 10 will be considered later.

No Triangular System Representation (DGPs 7, 8 and 12) DGP 7 allows for feedback from the dependent variables z_{1t} and z_{2t} to the variables z_{3t} and z_{4t}, but does not have any feedback between z_{1t} and z_{2t}. The bias is slightly higher than for the benchmark DGP 1, the same accounts for the variance of three of the estimated parameters. Consequently, the 90 percent band relative to the parameter values is marginally wider than for DGP 1. Skewness and kurtosis underlie only small changes.

For DGP 8 the regressors z_{3t} and z_{4t} are strongly exogenous for the first cointegrating relationship. As expected from the previous experiments, the strong exogeneity does not seem to have any influence on the performance of the Fully Modified OLS estimator; there do not appear significant differences between the estimates for the first and the second cointegrating relationship.

DGP 12 keeps z_{3t} and z_{4t} weakly exogenous for both cointegrating relationships but there is Granger causality from z_{1t} to z_{3t} and z_{4t}. Comparing the results of DGP 12 to those of DGP 9 again does not reveal any difference worth mentioning.

Concluding the results for the DGPs with exogenous regressors to the cointegrating relationships, it becomes clear that neither weak nor strong exogeneity seem to matter. This result should not be surprising because the Fully Modified OLS estimation procedure is a system estimator, which is asymptotically equivalent to full system maximum likelihood estimation and therefore asymptotically optimal, i.e. free of nuisance parameters. The lack of effect on the results of the Fully Modified OLS estimator by the validity of the condition of weak exogeneity is also reported for single–equation modeling [see Cappuccio and Lubian (1994)].

5.4 Autoregressive Error Processes with root 0.9 (DGP 5, 6, 10)

This section will analyze to which extent the Fully Modified OLS estimator is sensitive to autoregressive error processes with a root close to the unit root. Autoregressive error processes for one as well as for both cointegrating relationships will be considered.

Autoregressive Root 0.9 in First Cointegrating Vector (DGP 5) Modification of DGP 1 by the inclusion of a first order autoregressive error process in the first equation yields DGP 5. The change caused by this inclusion strikes the eye: Not only relative bias, but also standard deviation and 90 percent band show a significant deterioration. Interestingly to note hereby is the fact that the estimates for both cointegrating relationships are affected by the deterioration. Nevertheless, the degree of skewness is much smaller, but the distributions of the estimates are still skewed to the right; the changes in the kurtosis are minor.

Autoregressive Root 0.9 in Both Cointegrating Relationships (DGP 6, 10) As expected from the analysis above, DGPs 6 and 10, which add a first order autoregressive error process to both cointegrating relations in DGP 1 and 7 respectively, perform very badly in terms of relative bias, standard deviation, and 90 percent band. Whereas kurtosis remains almost constant, skewness is smaller when the autoregressive error is included. Note that these results confirm the single–equation analysis by Hargreaves (1994).

5.5 Moving Average Error Processes (DGPs 13 to 18)

The last issue, which should be covered by this Monte–Carlo analysis, is the question of the influence of moving average error processes on the Fully Modified OLS estimator. Five different DGPs will be considered. Amongst those are four with independent and two with positively and negatively interrelated moving average errors.

Independent Moving Average Errors (DGP 13, 16 to 18) DGP 13 considers a moving average parameter of 0.5 on the diagonal of the Θ matrix. Changes for the relative bias, standard deviation and 90 percent band behavior of the Fully Modified OLS estimator are barely noticeable. If there are movements at all, they seem to be more inclined to an improvement of the estimator. Kurtosis does not seem to be affected and skewness tends to be smaller (with one exception). This observation corroborates the single–equation analysis by Hargreaves (1994) apart from the fact that the kurtosis increased for his DGP.

The other DGPs have non–invertible (DGPs 16 and 17) or “nearly” non–invertible (DGP 18) moving average errors in the equations for the cointegrating relationships. Whereas the non–invertibility with root -1 does not seem to matter (DGP 16), the non–invertibility with root 1 (DGP 17) leads to very bad results for the relative bias, variance and the 90 percent band. Modifying the roots to 0.95 (DGP 18) ameliorates the results only minimally. Note that the last two DGPs are the only DGPs for which the kurtosis is positive, i.e. the Fully Modified OLS estimator has thicker tails than the Normal distribution.

Interrelated Moving Average Errors (DGP 14, 15) The matrix Θ is a full square matrix with 0.5 on the diagonal and 0.4 in all off–diagonal positions for DGP 14. DGP 15 modifies DGP 14 such that all the moving average coefficients are the negative of the previous Θ matrix. Surprisingly, the positively interrelated moving average errors have much less influence on the outcomes for the Fully Modified OLS estimator than the negatively interrelated errors. Whereas for relative bias, standard deviation and the 90 percent band, there is a hardly noticeable change for the worse for the positively interrelated errors, these measures deteriorate significantly for DGP 15. Kurtosis decreases slightly for both DGPs; this means that the tails of the distribution are getting fatter. Skewness decreases or remains the same for DGP 14 and increases significantly for three of the estimates for DGP 15. Again, these results are similar to those of Hargreaves.

5.6 Summary

After having surveyed the behavior of the Fully Modified OLS estimator for a variety of DGPs, a brief summary recalls the main results.

- Variance–covariance of the Error Process:
  A general statement regarding the reaction of the Fully Modified OLS estimator to changes in the variance–covariance matrix is not possible. Nevertheless, it should be noted that the quality of the Fully Modified OLS estimator seems to suffer mildly in the presence of intercorrelated errors.

- Exogeneity and Triangular System Representation:
  The performance of the Fully Modified OLS estimator does not seem to be influenced by the type of exogeneity of the regressors in the equations for the cointegrating relationships. This result is not unexpected, since the Fully Modified OLS estimator is a system

estimator. In exactly the same way, the possibility of representing the DGP in triangular system form with Gaussian errors does not appear to matter.

- Autoregressive Error Processes with Roots Close to Unity: Autoregressive error processes with a root close to unity seem to be the only experiments under consideration, which create a serious problem for the Fully Modified OLS estimator. This is simply due to the fact that the change of the A^1 matrix entails a significantly smaller signal–noise ratio for the cointegrated system.[20] Table 6 displays the roots of the signal–noise ratio matrix of the eighteen DGPs. Whereas for all other DGPs the smallest root is larger than 0.33, it hits the eye that for the DGPs with autoregressive processes with roots close to unity the smallest root is almost zero. Therefore, it is more difficult for the Fully Modified OLS estimator to distinguish the signal of the cointegrating space from the noise of the error process. As seen in the Monte–Carlo experiments, a sample size of 100 observations is not sufficient to guarantee reliable estimates of the cointegrating vectors. A significant extension of the sample size (at least by factor 5!), though, will eliminate the disadvantage of a small signal–noise ratio and lead to satisfying results of the Fully Modified OLS estimator.

Another way to understand the importance of the signal–noise ratio is to decompose the long–run static response matrix Π from equation (11) into the product of the adjustment coefficients α and cointegrating vectors β:[21]

$$\Pi = \underbrace{\begin{pmatrix} \Pi_{11} \\ \Pi_{21} \end{pmatrix}}_{=:\alpha} \underbrace{(\Pi_{11})^{-1} \begin{pmatrix} \Pi_{11} \Pi_{12} \end{pmatrix}}_{=:\beta}, \tag{18}$$

where Π is partioned in 2×2 matrices. If the matrix A^1 gets changed like in DGPs 5, 6, or 10, the cointegrating relationships β remain unaffected whereas the adjustment weights α become significantly

[20] Compare Kostial (1994).

[21] α and β are 4×2 matrices.

smaller. This leads to a smaller signal of the cointegrated system in relation to the noise of the error process and thus to problems for small sample sizes.

- Moving Average Error Processes:
  Moving average error processes do have a certain influence on the quality of the Fully Modified OLS estimator. The quantity of the influence depends on the type of interrelation: negatively interrelated moving average errors seem to cause more problems for the estimator than positively correlated errors.

In general, the distribution of the Fully Modified OLS estimator tends to be skewed to the right; a phenomenon, which might be explainable by the fact that for most of the DGPs the estimates lie to the left of the true value of the parameter. The kurtosis, though, is negative for almost all considered DGPs and parameters and thus the Fully Modified OLS estimator inclines to even thinner tails than the standard normal distribution.

6 The Role of the Rank of the Cointegrating Space

Since the rank of the cointegrating space was set to two for all Monte–Carlo experiments in the preceding section, this section is concerned with the question whether the rank of the cointegrating space plays a role for the behavior of the Fully Modified OLS estimator. The underlying idea is simply that the higher is the rank of the cointegrating space in relation to the dimension of the system, the closer is the system to an $I(0)$ system and thereby easier to estimate.

Three DGPs with very similar structures are considered to demonstrate the impact of the rank of the cointegrating space on the precision of the Fully Modified OLS estimator. Again DGP 1 with two cointegrating relationships is the benchmark DGP (denoted in this section as M_2) and the other DGPs M_1 (with one cointegrating relationship) and M_3

(with three cointegrating vectors) are generated by manipulating the autoregressive error process in M_1 (see table 7). Monte–Carlo experiments were run for 40, 60 and 100 observations.

Tables 8 and 9 display bias and variance, and the 5 and 95 percent quartiles.[22] The cointegrating relationships follow for M_1 as $(1\ c_{11}\ c_{12}\ c_{13})$, for M_2 as

$$\begin{pmatrix} 1 & 0 & c_{11} & c_{12} \\ 0 & 1 & c_{21} & c_{22} \end{pmatrix},$$

and for M_3 as

$$\begin{pmatrix} 1 & 0 & 0 & c_{11} \\ 0 & 1 & 0 & c_{21} \\ 0 & 0 & 1 & c_{31} \end{pmatrix}.$$

The tables illustrate impressively that the rank of the cointegrating space is a significant element for the judgement of the quality of the Fully Modified OLS estimator. Whereas the results for bias, variance and 90 percent band for M_1 are not convincing even for a high number of observations, M_3 is estimated well for only 40 observations. The general rule is that the higher the rank of the cointegrating space (in relation to the dimension of the system), the higher is the precision of the Fully Modified OLS estimator (in terms of narrower 90 percent bands, smaller bias and variance). Skewness and kurtosis (which are not reported here), though, do not seem to underlie a general pattern.

[22]The normed cointegrating vector for M_1 is $(1\ -0.2\ 1.4\ -0.4)$; for M_2 the normed cointegrating vectors are

$$\begin{pmatrix} 1 & 0 & 1 & 0.03 \\ 0 & 1 & 1 & 0.38 \end{pmatrix},$$

and for M_3

$$\begin{pmatrix} 1 & 0 & 0 & -0.81 \\ 0 & 1 & 0 & 0.33 \\ 0 & 0 & 1 & 0.05 \end{pmatrix}.$$

7 Permutations as a Means to Determine the Rank of the Cointegrating Space

This section owes its existence the fact that many systems, which are set up to estimate cointegrating relationships, include variables, which appear to be stationary at least over some time periods. In the case of the inclusion of a nearly stationary variable into a system, though, there exist permutations of the variables such that the system cannot be represented in triangular system form for some periods of the sample size. An estimation of this special ordering of the variables will give misleading results.

In order to avoid this problem with the triangular system representation, the Fully Modified OLS estimator should not only be applied to the original data set, but also to transformations of the data set by permutation matrices. Under the correct hypothesis on the rank of the cointegrating space, estimation results derived by the Fully Modified OLS estimator should be invariant under different permutations or — in case of differing outcomes — be explainable as a product of the mis–specification mentioned above.

Since the application of the procedure described above together with an underestimation of the rank of the cointegrating space yields the pleasant byproduct of estimates of different vectors in the cointegrating space, the motivation for this section was born. This section wants to introduce the "permutation procedure" as a means determining the number of cointegrating vectors as an alternative to the rank tests.

For that purpose, benchmark DGP 1 is estimated under the assumption that the cointegrating space has rank one, two or three. The Monte–Carlo experiments are on the original data set, and the permutations to $(z_{2t}, z_{3t}, z_{1t}, z_{4t})$ and $(z_{3t}, z_{4t}, z_{1t}, z_{2t})$.[23] Table 10 reports mean and variance of the permutated Fully Modified OLS estimates.

[23] For illustrative purposes, the other three existing permutations of the sets of driving and non–driving variables will not be reported.

To begin with, the estimates derived under the assumption of a one–dimensional cointegrating space are considered. In order to get comparable results for the different permutations, the means were normed on the first variable z_{1t}. The cointegrating vector for the original data set is mean–estimated as

$$\beta_{1234} = (1 \quad -0.466 \quad 0.536 \quad -0.212)\,,$$

for the 2314–permutated data set the estimate is

$$\beta_{2314} = (1 \quad -2.198 \quad -1.174 \quad -0.864)\,,$$

and for the permutation 3412 it is

$$\beta_{3412} = (1 \quad 1.000 \quad 2.008 \quad 0.343)\,.$$

These values for the Monte–Carlo mean give clear evidence that the rank of the cointegrating space is underestimated. It should be noted that, though the variance is not very small (and similarly, as expected, the band formed by the 5 and 95 percent quartiles [which are not reported]), the Monte–Carlo mean is very precise.

Turning to the results for the assumption of a two–dimensional cointegrating space and norming on z_{1t} and z_{2t}, the estimated cointegrating space for the original data set is spanned by

$$\beta_{1234} = \begin{pmatrix} 1 & 0 & 0.991 & -0.036 \\ 0 & 1 & 0.985 & 0.377 \end{pmatrix}.$$

For permutation 2314 the estimated cointegrating space is given by

$$\beta_{2314} = \begin{pmatrix} 1 & 0 & 1.008 & -0.034 \\ 0 & 1 & 0.998 & 0.377 \end{pmatrix},$$

and for permutation 3412 it is

$$\beta_{3412} = \begin{pmatrix} 1 & 0 & 1.002 & -0.035 \\ 0 & 1 & 1.003 & 0.379 \end{pmatrix}.$$

In other words, the estimates for the cointegrating vectors are consistent under the permutations.

The question, which appears now is whether, regardless of the fact of consistency under the (correct) assumption of the cointegrating space having rank two, the estimates derived under the assumption of a three-dimensional cointegrating space are still consistent under the permutations. Normation on z_{1t}, z_{2t} and z_{3t} gives the estimated cointegrating space for the original data set as

$$\beta_{1234} = \begin{pmatrix} 1 & 0 & 0 & 0.275 \\ 0 & 1 & 0 & 0.698 \\ 0 & 0 & 1 & -0.313 \end{pmatrix},$$

for permutation 2314 as

$$\beta_{2314} = \begin{pmatrix} 1 & 0 & 0 & 0.155 \\ 0 & 1 & 0 & 0.259 \\ 0 & 0 & 1 & 0.119 \end{pmatrix},$$

and for permutation 3412 as

$$\beta_{3412} = \begin{pmatrix} 1 & 0 & 0 & -0.398 \\ 0 & 1 & 0 & 0.021 \\ 0 & 0 & 1 & 0.359 \end{pmatrix}.$$

These results exhibit convincingly that the assumption of a three-dimensional cointegrating space was wrong and that accordingly the rank of the cointegrating space of the DGP has to be two.[24] Or, expressed differently, permutations of the data set can be a very helpful mean to determine the rank of the cointegrating space.

8 Application to the Data Set of Clements and Mizon (1991)

This section is the crowning of the preceding series of Monte-Carlo experiments, since it illustrates the usefulness of the Fully Modified OLS

[24]Note that results derived by using different kernels (not reported in the table) can lead to different estimates of the cointegrating space if the rank of the cointegrating space is underestimated.

estimator as a system estimator for a particular data set. The data set, which will be re-estimated with the Fully Modified OLS estimator, is the data used by Clements and Mizon, who estimated the cointegrating relationships with the Johansen procedure. They analyze the determination of earnings, prices, productivity, hours worked, and unemployment in the UK. The Johansen tests on the rank of the cointegrating rank confirm the existence of at least three cointegrating vectors and a fourth cointegrating vector is rejected at the 5% significance level.

Since in the case of three or four cointegrating vectors in a five dimensional system the rank of the cointegrating space in relation to the dimension of the system is quite high and, moreover, the number of observations is almost hundred, the Monte–Carlo experiments assure that relatively precise estimates can be expected. The same dummies as proposed by Clements and Mizon are partialled out and then the prewhitened Fully Modified OLS estimator with the Bartlett kernel and an automatic bandwidth parameter is applied to the transformed system.

Table 11 displays the estimates derived with the Fully Modified OLS estimator under the assumption of the cointegrating space having rank three and four. Since the number of permutations of a five dimensional set of variables is 5!, the analysis was confined to permutations between the set of driving and non–driving variables, which is $\binom{5}{3}$ in the case of three and $\binom{5}{4}$ in the case of four cointegrating relationships. The permutation "45123", for example, is equivalent to the ordering $z_{3t},\ z_{4t},\ z_{5t},\ z_{1t},\ z_{2t}$.[25]

[25]The cointegrating relationships follow from the table as

$$\begin{pmatrix} 1 & 0 & 0 & c_{11} & c_{12} \\ 0 & 1 & 0 & c_{21} & c_{22} \\ 0 & 0 & 1 & c_{31} & c_{32} \end{pmatrix}$$

for the assumption of three cointegrating relationships and as

$$\begin{pmatrix} 1 & 0 & 0 & 0 & c11 \\ 0 & 1 & 0 & 0 & c12 \\ 0 & 0 & 1 & 0 & c13 \\ 0 & 0 & 0 & 1 & c14 \end{pmatrix}$$

for the case of a four dimensional cointegrating space.

Under the assumption of a three dimensional cointegrating space, the estimates are clearly inconsistent under the permutations, whereas the assumption of the cointegrating space having rank four yields consistent estimates for all but one permutation. The inconsistency of permutation "15234", though, is a logical consequence of the ordering of the variables: The inflation rate, which proves to be stationary, is chosen as the driving variable of the system and thus the system is mis–specified. Thus, the re-estimation with the Fully Modified OLS estimator leads to the conclusion that the rank of the cointegrating space is four.

After having determined the cointegrating vectors with the Fully Modified OLS estimator, these estimates should be compared with the estimates derived by Clements and Mizon (1991). They interpret the four dimensional cointegrating space as two stationary variables, inflation rate and average hours of work, plus two linear combinations of the variables given by:

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 0.06 \\ 0 & 0 & 1 & 0 & -0.14 \end{pmatrix}. \tag{19}$$

The first row is a target relation for the real average earnings adjusted for output per hour worked, for which the logged unemployment rate exerts a negative influence on the evolution of productivity adjusted real earnings. The second row indicates a complementarity between the average productivity and the unemployment rate — a feature of recent UK history.

The cointegrating vectors from equation (19) and the stationarity of the inflation rate can be confirmed by the Fully Modified OLS estimator. The stationarity of the hours worked, though, exhibits a small but strong relationship with the unemployment rate for the Fully Modified OLS estimator. This is seen especially by the estimates for different permutations under the assumption of the existence of three cointegrating relationships. Note that this observation is not due to a disagreement of the different estimators, but to a different interpretation of the results.[26]

[26] Tests on the different "interpretations" do not lead to rejections for both cases.

9 Conclusion

This paper provided an extensive Monte–Carlo study of the performance of the Fully Modified OLS estimator. In a first step, the "optimal" Fully Modified OLS estimator was chosen by considering various alternatives for estimating the long–run variance–covariance matrix. Included in these alternatives are the use of different kernels and different bandwidths as well as the procedure of prewhitening as an alternative to directly inserting the estimated error process into the formula for the long–run variance–covariance matrix. The best combination of the three alternatives appears to be the prewhitened Fully Modified OLS estimator with the Bartlett or the truncated kernel and an automatic bandwidth choice.

Section 5 analyzed the reaction of the Fully Modified OLS estimator to non–Gaussian error processes and different kinds of exogeneity. The estimator proves to be most sensitive to autoregressive and moving average errors with one root close to the unit root. For the autoregressive case, this behavior can be explained theoretically by the reduction of the signal–noise ratio of the cointegrated system through the inclusion of the autoregressive error. The influence of variance–covariance matrices of the error process different from the identity matrix, the type of exogeneity as well as the possibility of representing the DGP in triangular system form with Gaussian errors do not seem to play a significant role when estimating with the Fully Modified OLS estimator.

The messages of the following sections are of relevance for applied work with the Fully Modified OLS estimator. Section 6 illustrates that the higher the rank of the cointegrating space, the more precise will be the estimates. Or, in other words, the closer the $I(1)$ system is to a stationary system, the easier it is to estimate. The section next to the last highlights the possibility of exploiting permutations of the data set to determine the rank of the cointegrating space. System estimates derived with the Fully Modified OLS estimator should only be accepted if they are invariant under different permutations. Applying the "permutation procedure" to the data set, which has been estimated by Clements and

Mizon (1991) with the Johansen procedure, helps to determine the rank of the cointegrating space and confirms the estimated results.

Table 1: DGPs

M	$B_{.,1:2}$		$A1_{.,1:2}$		Θ				Ω			
1	1	−2/7	0	0	0	0	0	0	1	0	0	0
	1	3/4	0	0	0	0	0	0	0	1	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1
2	1	−2/7	0	0	0	0	0	0	1	0	0	0
	1	3/4	0	0	0	0	0	0	0	25	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1
3	1	−2/7	0	0	0	0	0	0	1	0.8	0.8^2	0.8^3
	1	3/4	0	0	0	0	0	0	0.8	1	0.8	0.8^2
	1	4/3	0	0	0	0	0	0	0.8^2	0.8	1	0.8
	1	5	0	0	0	0	0	0	0.8^3	0.8^2	0.8	1
4	1	−2/7	0	0	0	0	0	0	1	−0.8	0.8^2	-0.8^3
	1	3/4	0	0	0	0	0	0	−0.8	1	−0.8	0.8^2
	1	4/3	0	0	0	0	0	0	0.8^2	−0.8	1	−0.8
	1	5	0	0	0	0	0	0	-0.8^3	0.8^2	−0.8	1
5	1	−2/7	0.9	0	0	0	0	0	1	0	0	0
	1	3/4	0	0	0	0	0	0	0	1	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1

Continuation of table 1: DGPs

M	$B_{.,1:2}$		$A1_{.,1:2}$		Θ				Ω			
6	1	−2/7	0.9	0	0	0	0	0	1	0	0	0
	1	3/4	0	0.9	0	0	0	0	0	1	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1
7	1	0	0	0	0	0	0	0	1	0	0	0
	0	3/4	0	0	0	0	0	0	0	1	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1
8	1	−2/7	0	0	0	0	0	0	1	0	0	0
	0	3/4	0	0	0	0	0	0	0	1	0	0
	0	4/3	0	0	0	0	0	0	0	0	1	0
	0	5	0	0	0	0	0	0	0	0	0	1
9	1	0	0	0	0	0	0	0	1	0	0	0
	0	3/4	0	0	0	0	0	0	0	1	0	0
	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	1
10	1	0	0.9	0	0	0	0	0	1	0	0	0
	0	3/4	0	0.9	0	0	0	0	0	1	0	0
	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	1

Continuation of table 1: DGPs

M	$B_{.,1:2}$		$A1_{.,1:2}$		Θ				Ω			
11	1	0	0	0	0	0	0	0	1.17	0.62	0.45	−0.45
	0	3/4	0	0	0	0	0	0	0.62	1.22	0.58	−0.58
	0	0	0	0	0	0	0	0	0.45	0.58	1.17	0.27
	0	0	0	0	0	0	0	0	−0.45	−0.58	0.27	1.17
12*	1	−2/7	0	0	0	0	0	0	1	0	0	0
	1	3/4	0	0	0	0	0	0	0	1	0	0
	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	1
13	1	−2/7	0	0	0.5	0	0	0	1	0	0	0
	1	3/4	0	0	0	0.5	0	0	0	1	0	0
	1	4/3	0	0	0	0	0.5	0	0	0	1	0
	1	5	0	0	0	0	0	0.5	0	0	0	1
14	1	−2/7	0	0	0.5	0.4	0.4	0.4	1	0	0	0
	1	3/4	0	0	0.4	0.5	0.4	0.4	0	1	0	0
	1	4/3	0	0	0.4	0.4	0.5	0.4	0	0	1	0
	1	5	0	0	0.4	0.4	0.4	0.5	0	0	0	1
15	1	−2/7	0	0	−0.5	−0.4	−0.4	−0.4	1	0	0	0
	1	3/4	0	0	−0.4	−0.5	−0.4	−0.4	0	1	0	0
	1	4/3	0	0	−0.4	−0.4	−0.5	−0.4	0	0	1	0
	1	5	0	0	−0.4	−0.4	−0.4	−0.5	0	0	0	1

*: Granger causality from z_{1t} on z_{3t} with the coefficient0.5 and on z_{4t} with the coefficient -0.5.

Continuation of table 1: DGPs

M	$B_{.,1:2}$		$A1_{.,1:2}$		Θ				Ω			
16	1	−2/7	0	0	1	0	0	0	1	0	0	0
	1	3/4	0	0	0	1	0	0	0	1	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1
17	1	−2/7	0	0	−1	0	0	0	1	0	0	0
	1	3/4	0	0	0	−1	0	0	0	1	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1
18	1	−2/7	0	0	−0.95	0	0	0	1	0	0	0
	1	3/4	0	0	0	−0.95	0	0	0	1	0	0
	1	4/3	0	0	0	0	0	0	0	0	1	0
	1	5	0	0	0	0	0	0	0	0	0	1

Table 2:

Automatic vs. Fixed Bandwidth: DGP 1

		aut.b.	fixed bandwidth		
			$S = 1$	$S = 3$	$S = 6$
bias	c_{11}	0.009	0.003	0.005	0.009
	c_{12}	0.002	0.002	0.002	0.002
	c_{21}	0.015	0.007	0.009	0.015
	c_{22}	0.002	0.001	0.001	0.002
var.[1]	c_{11}	5.81	5.45	5.28	5.57
	c_{12}	0.95	0.95	0.92	0.96
	c_{21}	4.32	3.20	3.19	4.23
	c_{22}	0.76	0.55	0.57	0.74
q05	c_{12}	1.048	1.054	1.049	1.046
	c_{12}	-0.027	-0.027	-0.027	-0.029
	c_{21}	1.012	1.013	1.011	1.009
	c_{22}	0.382	0.381	0.381	0.382
q95	c_{11}	0.918	0.934	0.948	0.929
	c_{12}	-0.048	-0.048	-0.047	-0.051
	c_{21}	0.925	0.953	0.958	0.947
	c_{22}	0.367	0.371	0.372	0.365
s	c_{11}	0.25	0.50	0.26	0.63
	c_{12}	0.59	0.52	0.74	1.77
	c_{21}	3.50	2.06	3.28	-12.13
	c_{22}	2.81	2.03	1.65	2.76
k	c_{11}	-0.13	-0.14	-0.08	-0.18
	c_{12}	-0.14	-0.12	-0.20	-0.26
	c_{21}	-0.14	-0.24	-0.29	-0.46
	c_{22}	-0.22	-0.24	-0.32	-0.22

[1] var. = variance $\times 10^4$.

Continuation table 2:
Automatic vs. Fixed Bandwidth: DGP 6

		aut.b.	fixed bandwidth		
			$S = 1$	$S = 3$	$S = 6$
bias	c_{11}	0.056	0.054	0.050	0.050
	c_{12}	0.013	0.012	0.013	0.012
	c_{21}	0.017	0.015	0.043	0.050
	c_{22}	0.006	0.007	0.008	0.007
var.[1]	c_{11}	57.77	53.22	54.09	53.50
	c_{12}	12.90	11.56	11.65	11.58
	c_{21}	74.10	28.59	26.23	26.61
	c_{22}	11.28	4.98	4.75	4.67
q05	c_{12}	1.516	1.478	1.472	1.497
	c_{12}	0.071	0.069	0.086	0.091
	c_{21}	1.209	1.226	1.141	1.171
	c_{22}	0.416	0.416	0.415	0.426
q95	c_{11}	0.373	0.370	0.375	0.420
	c_{12}	-0.181	-0.181	-0.173	-0.161
	c_{21}	0.687	0.699	0.693	0.652
	c_{22}	0.316	0.320	0.329	0.329
s	c_{11}	0.08	3.46	2.74	2.22
	c_{12}	0.25	4.36	3.07	3.72
	c_{21}	0.24	4.91	4.05	80.43
	c_{22}	0.30	4.64	-10.33	-7.08
k	c_{11}	-0.18	-0.13	-0.09	-0.14
	c_{12}	-0.22	-0.16	-3.33	-4.74
	c_{21}	-0.18	-0.18	0.04	-0.04
	c_{22}	-0.23	-0.21	-0.21	-0.31

[1] var. = variance $\times 10^4$.

Table 3: Relative Bias and Variance for DGP 1 to 18

DGP	bias				variance $\times 10^4$			
	c_{11}	c_{12}	c_{21}	c_{22}	c_{11}	c_{12}	c_{21}	c_{22}
1	0.009	-0.053	0.015	0.007	5.81	27.48	4.32	2.52
2	0.004	0.013	0.018	0.005	2.87	17.14	9.29	4.96
3	0.043	0.027	0.054	-0.002	9.60	32.47	11.47	3.49
4	0.011	0.124	0.000	0.003	7.99	71.76	1.39	1.31
5	0.005	-0.530	0.038	0.075	31.76	137.13	38.20	15.31
6	0.056	0.359	0.017	0.017	57.77	374.18	74.10	29.75
7	0.038	0.013	0.039	0.016	13.67	4.55	8.77	3.92
8	0.005	0.108	0.056	0.027	10.02	22.51	9.78	5.18
9	0.013	0.043	0.014	0.068	6.61	24.09	3.41	18.27
10	0.010	0.046	0.143	0.380	68.76	225.23	30.67	150.48
11	0.017	0.046	0.017	0.074	5.86	22.64	3.61	19.63
12	0.019	0.300	0.010	0.011	5.69	85.38	4.99	7.05
13	0.002	0.034	0.011	0.004	5.47	26.40	3.42	1.50
14	0.017	0.046	0.023	0.000	6.32	23.21	6.02	1.96
15	0.055	0.357	0.065	0.040	10.57	63.25	12.07	6.88
16	0.001	0.063	0.007	0.002	4.59	30.72	2.72	1.47
17	0.189	0.732	0.267	0.090	20.09	755.28	27.95	9.45
18	0.187	0.670	0.263	0.088	19.80	71.92	27.54	9.36

Table 4: q05 and q95 for DGPs 1 to 18

DGP	q05				q95			
	c_{11}	c_{12}	c_{21}	c_{22}	c_{11}	c_{12}	c_{21}	c_{22}
1	1.048	-0.027	1.012	0.382	0.918	-0.048	0.925	0.367
2	1.026	-0.027	1.068	0.398	0.960	-0.041	0.859	0.354
3	1.019	-0.020	1.014	0.395	0.825	-0.045	0.789	0.366
4	1.063	-0.015	1.014	0.382	0.891	-0.069	0.984	0.372
5	1.330	0.040	1.391	0.409	0.626	-0.066	0.527	0.278
6	1.516	0.071	1.209	0.416	0.373	-0.181	0.687	0.316
7	0.778	-0.137	2.398	0.339	0.555	-0.153	1.957	0.310
8	1.407	-0.045	2.349	0.336	1.106	-0.068	1.872	0.299
9	0.752	-0.115	2.360	0.354	0.646	-0.191	2.194	0.227
10	1.149	0.199	2.636	0.695	0.082	-0.528	1.081	-0.407
11	0.740	-0.119	2.346	0.350	0.650	-0.191	2.176	0.218
12	1.032	-0.019	1.040	0.403	0.904	-0.083	0.928	0.344
13	1.057	-0.026	1.013	0.382	0.933	-0.047	0.942	0.370
14	1.037	-0.025	1.014	0.387	0.900	-0.043	0.894	0.372
15	1.025	-0.031	1.014	0.381	0.798	-0.078	0.757	0.326
16	1.051	-0.025	1.012	0.383	0.949	-0.048	0.959	0.372
17	0.989	-0.035	0.992	0.378	0.549	-0.090	0.375	0.299
18	0.991	-0.035	0.993	0.379	0.553	-0.090	0.386	0.300

Table 5: Skewness and Kurtosis for DGP 1 to 18

DGP	skewness				kurtosis			
	c_{11}	c_{12}	c_{21}	c_{22}	c_{11}	c_{12}	c_{21}	c_{22}
1	0.50	0.59	3.50	2.81	-0.13	-0.14	-0.14	-0.22
2	0.20	0.01	0.80	0.21	-0.11	-0.13	-0.14	-0.09
3	2.92	-0.20	3.86	-0.16	-0.17	-0.16	-0.16	-0.18
4	0.61	0.54	0.33	1.45	-0.18	-0.17	-0.16	-0.23
5	0.12	-0.17	0.05	0.38	-0.14	-0.13	-0.14	-0.10
6	0.08	0.25	0.24	0.30	-0.18	-0.22	-0.18	-0.23
7	0.83	0.65	3.30	2.64	-0.09	-0.11	-0.09	-0.10
8	2.24	2.36	3.78	3.72	-0.06	-0.04	-0.09	-0.08
9	0.45	0.50	1.93	2.31	-0.13	-0.12	-0.17	-0.13
10	0.21	0.07	0.73	0.47	-0.14	-0.16	-0.09	-0.12
11	0.83	0.58	3.85	3.29	-0.13	-0.13	-0.13	-0.13
12	1.05	1.07	0.51	0.40	-0.10	-0.10	-0.14	-0.13
13	0.27	0.35	2.38	1.50	-0.12	-0.09	-0.16	-0.17
14	1.10	0.07	3.56	-0.03	-0.15	-0.12	-0.15	-0.17
15	2.16	2.53	3.14	3.70	-0.07	-0.08	-0.08	-0.04
16	0.06	0.23	1.74	0.76	-0.11	-0.14	-0.20	-0.24
17	1.20	1.04	0.89	0.86	0.08	0.04	0.10	0.08
18	1.00	0.98	0.86	0.88	0.07	0.07	0.09	0.08

Table 6: Signal–Noise Ratio of DGP 1 to 18

DGP	roots		DGP	roots		DGP	roots	
1	0.70	0.33	7	0.93	0.35	13	0.71	0.33
2	0.81	0.34	8	0.93	0.34	14	0.74	0.33
3	0.81	0.39	9	0.44	0.34	15	0.75	0.33
4	0.74	0.33	10	0.18	0.05	16	0.76	0.33
5	0.71	0.05	11	0.45	0.35	17	0.78	0.34
6	0.61	0.05	12	0.78	0.32	18	0.76	0.34

Table 11: Clements and Mizon (1991): FM–OLS

permutations	estimates under the assumption of rank 3					
	c_{11}	c_{12}	c_{21}	c_{22}	c_{31}	$c32$
12345	3.90	-0.11	0.62	0.02	3.40	-0.15
12435	16.96	0.26	1.61	0.05	14.15	0.15
14235	0.18	-0.23	0.46	0.01	-0.21	-0.27
14235	6.12	-0.03	1.18	0.04	5.18	-0.09
14253	9.96	0.09	1.50	0.05	8.67	0.02
14523	7.37	-0.01	1.43	0.04	5.83	-0.08
41253	14.45	0.17	1.43	0.04	11.62	0.06
41253	-1.32	-0.33	0.33	0.01	-1.26	-0.35
41523	28.46	0.48	2.36	0.06	21.13	0.27
45123	11.07	0.02	1.73	0.04	9.84	-0.04
	estimates under the assumpion of rank 4					
	c_{11}	c_{12}	c_{13}	$c14$		
12345	-0.21	0.00	-0.24	0.03		
12354	-0.28	-0.01	-0.30	0.03		
12534	-0.22	0.00	-0.24	0.03		
15234	-6.02	-0.78	-5.36	0.50		
51234	-0.28	-0.00	-0.30	0.03		

Table 7: Error Processes for M_1, M_2, M_3

M_1	M_2	M_3
$u_{1t} = \epsilon_{1t}$	$u_{1t} = \epsilon_{1t}$	$u_{1t} = \epsilon_{1t}$
$u_{2t} = u_{2t-1} + \epsilon_{2t}$	$u_{2t} = \epsilon_{2t}$	$u_{2t} = \epsilon_{2t}$
$u_{3t} = 1.5u_{3t-1} - 0.5u_{3t-2} + \epsilon_{3t}$	$u_{3t} = 1.5u_{3t-1} - 0.5u_{3t-2} + \epsilon_{3t}$	$u_{3t} = \epsilon_{3t}$
$u_{4t} = u_{4t-1} + \epsilon_{4t}$	$u_{4t} = u_{4t-1} + \epsilon_{4t}$	$u_{4t} = u_{4t-1} + \epsilon_{4t}$

Table 8: Rank of the Cointegrating Space: Bias and Variance

r	obser.	bias				variance $\times 10^4$			
1		c_{11}	c_{12}	c_{13}		c_{11}	c_{12}	c_{13}	
	40	-0.059	0.250	-0.166		7.78	30.48	31.01	
	60	-0.040	0.164	-0.115		5.62	21.52	22.08	
	100	-0.023	0.090	-0.067		3.51	13.11	13.21	
2		c_{11}	c_{12}	c_{21}	c_{22}	c_{11}	c_{12}	c_{21}	c_{22}
	40	0.043	0.007	0.065	0.007	14.93	2.52	13.32	2.36
	60	0.022	0.004	0.030	0.003	9.67	1.70	8.03	1.53
	100	0.009	0.002	0.015	0.000	5.81	0.95	4.32	0.76
3		c_{11}	c_{21}	c_{31}		c_{11}	c_{21}	c_{31}	
	40	0.000	0.001	0.001		0.93	0.86	0.90	
	60	0.000	0.000	0.000		0.64	0.49	0.57	
	100	0.000	0.000	0.000		0.38	0.28	0.34	

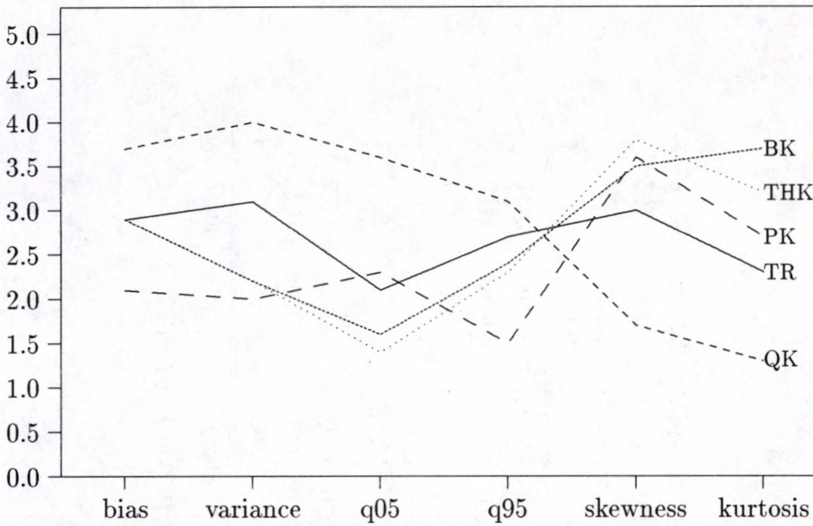
Table 9: Rank of the Cointegrating Space: q05 and q95

r	obs.	q05				q95			
1		c_{11}	c_{12}	c_{13}		c_{11}	c_{12}	c_{13}	
	40	-0.038	1.417	0.173		-0.211	0.739	-0.544	
	60	-0.079	1.411	0.019		-0.207	0.934	0.486	
	100	-0.128	1.405	-0.151		-0.203	1.129	-0.438	
2		c_{11}	c_{12}	c_{21}	c_{22}	c_{11}	c_{12}	c_{21}	c_{22}
	40	1.094	-0.016	1.046	0.389	0.757	-0.074	0.760	0.337
	60	1.069	-0.022	1.026	0.384	0.857	-0.060	0.859	0.353
	100	1.048	-0.027	1.012	0.382	0.918	-0.048	0.925	0.367
3		c_{11}	c_{21}	c_{31}		c_{11}	c_{21}	c_{31}	
	40	-0.071	0.339	0.054		-0.091	0.322	0.035	
	60	-0.074	0.337	0.051		-0.088	0.327	0.039	
	100	-0.076	0.335	0.050		-0.085	0.330	0.042	

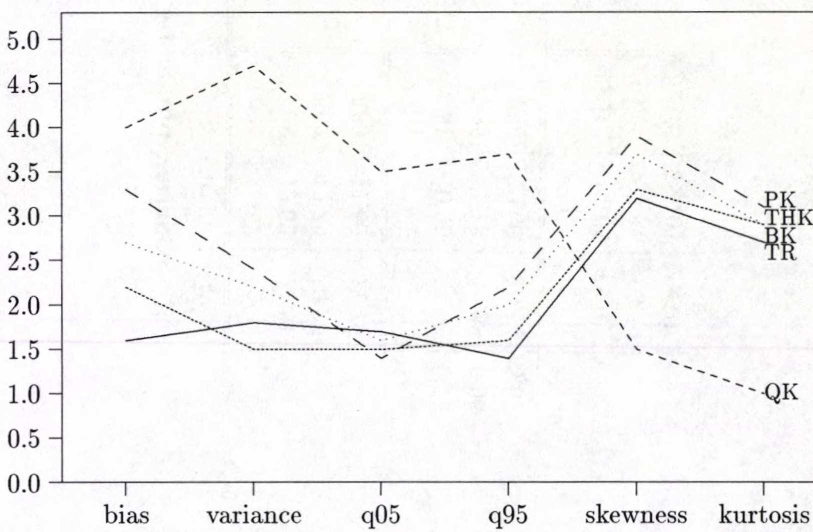
Table 10: Permutations: Mean and Variance

r	perm	mean				variance $\times 10^4$			
1		c_{11}	c_{12}	c_{13}		c_{11}	c_{12}	c_{13}	
	1234	-0.466	0.536	-0.212		12.98	12.62	4.81	
	2314	0.534	-0.455	0.393		12.92	12.89	0.87	
	3412	0.171	0.498	0.498		4.83	11.60	11.70	
2		c_{11}	c_{12}	c_{21}	c_{22}	c_{11}	c_{12}	c_{21}	c_{22}
	1234	0.991	-0.036	0.985	0.377	5.81	0.95	4.32	0.76
	2314	-0.990	0.411	0.992	-0.034	5.88	1.18	4.98	0.96
	3412	0.914	0.084	-2.419	2.416	3.86	2.06	10.54	5.57
3		c_{11}	c_{21}	c_{31}		c_{11}	c_{21}	c_{31}	
	1234	0.275	0.698	-0.313		4×10^4	4×10^4	4×10^4	
	2314	0.259	0.119	0.155		94.30	93.84	93.39	
	3412	-16.853	46.963	18.714		2×10^5	5×10^5	2×10^5	

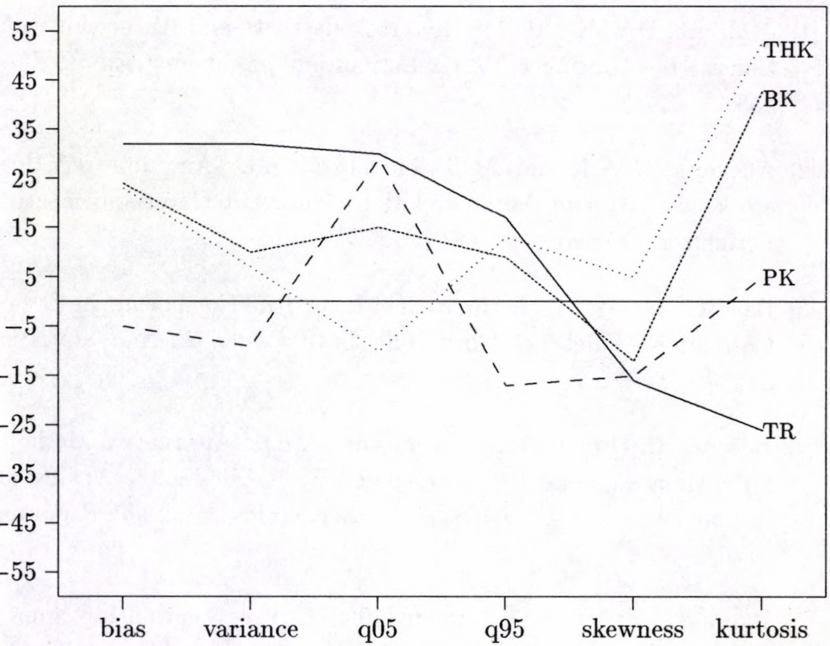
Graph 1: Appropriate Kernel for Non–Prewhitened Estimator



Graph 2: Appropriate Kernel for Prewhitened Estimator



Graph 3: Prewhitening versus Non–Prewhitening



References

[1] Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica 59, 817–858.*

[2] Andrews, D.W.K and J.C. Monahan (1992), "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator", *Econometrica 60, 953–966.*

[3] Bårdsen, G. (1989), "Estimation of Long–Run Coefficients in Error–Correction Models", *Oxford Bulletin of Economics and Statistics 345–350.*

[4] Bewley, R., Orden, D. and L. Fisher (1991), "Alternative Methods for Estimating Long–Run Responses with Application to Australian Import Demand", *UNSW School of Economics Discussion Paper no. 91/6.*

[5] Cappuccio, N. and D. Lubian (1992), "The Relationship among Some Estimators of the Cointegrating Coefficients", *Manuscript, Department of Economics, University of Padova.*

[6] Cappuccio, N. and D. Lubian (1994), "Fully Modified Estimation of Cointegrating Vectors Via VAR Prewhitening: A Simulation Study" *Manuscript, Department of Economics, University of Padova, presented at ESEM 94.*

[7] Clements, M.P. and G.E. Mizon (1991), "Empirical Analysis of Macroeconomic Time Series", *European Economic Review 36, 887–932.*

[8] Dunsmuir, W. (1979), "A Central Limit Theorem for Parameter Estimation in Stationary Vector Time Series and Its Application to Models for a Signal Observed with Noise", *Annals of Statistics 7, 490–506.*

[9] Dunsmuir, W. and E.J. Hannan (1976), "Vector Linear Time Series Models", *Advances in Applied Probability 8, 339–364.*

[10] Engle, R.F. and C.W.J. Granger (1987), "Co–integration and Error Correction: Representation, Estimation, and Testing", *Econometrica 55 (2), 251–276.*

[11] Engle, R.F. and S. Yoo (1991), "Cointegrated Economic Time Series: An Overview with New Results, chapter 12 in: *Engle, R.F. and C.W.J. Granger (eds.), "Long–Run Economic Relationships; Readings in Cointegration, Oxford University Press.*

[12] Hansen, B.E. (1992), "Tests for Parameter Instability in Linear Models", *Journal of Business and Economic Statistics 10, 321–335.*

[13] Hansen, B.E. and P.C.B. Phillips (1990), "Estimation and Inference in Models of Cointegration: A Simulation Study", *Advances in Econometrics 8, JAI, 225–248.*

[14] Hargreaves, C. (1994), "A Review of Methods of Estimating Cointegrating Relationships", in *Hargreaves (ed.), "Nonstationary Time Series Analyses and Cointegration"*, *Oxford University Press.*

[15] Inder, B. (1991), "Estimating Long Run Relationships in Economics: A Comparison of Different Approaches", *mimeo, Department of Econometrics, Monash University Australia.*

[16] Inder, B. (1993), "Finite Sample Arguments for Appropriate Estimation of Cointegrating Relationships", *mimeo, Department of Econometrics, Monash University Australia.*

[17] Johansen, S. (1988), "Statistical Analysis of Cointegrating Vectors", *Journal of Economic Dynamics and Control 12, 231–254.*

[18] Johansen (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica 59, 1551–1580.*

[19] Kostial, K (1994), "The Role of the Signal–Noise Ratio in Cointegrated Systems", *EUI working paper no.94/27.*

[20] Park, J.Y. (1992), "Canonical Cointegrating Regressions", *Econometrica 60, 119–143.*

[21] Phillips, P.C.B. (1991a), "Optimal Inference in Cointegrated Systems", *Econometrica 59, 283–306.*

[22] Phillips, P.C.B. (1991b), "Spectral Regression for Cointegrated Time Series", in: *Barnett, W., Powell, J. and G. Tauchen (eds.), "Nonparametric and Semiparametric Methods in Economics and Statistics, New York: Cambridge University Press.*

[23] Phillips, P.C.B. and B.E. Hansen (1990), "Statistical Inference in Instrumental Variables Regression with $I(1)$ Processes", *Review of Economic Studies 57, 99–125.*

[24] Phillips, P.C.B. and M. Loretan (1991), "Estimating Long–Run Economic Equilibria", *Review of Economic Studies 58, 407–436.*

[25] Saikonnen, P. (1991), "Asymptotically Efficient Estimation of Cointegration Regressions", *Econometric Theory 7, 1–21.*

[26] Stock, J.H. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors", *Econometrica 55, 1035–1056.*

EUI WORKING PAPERS

EUI Working Papers are published and distributed by the
European University Institute, Florence

Copies can be obtained free of charge
– depending on the availability of stocks – from:

The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
Italy

Please use order form overleaf

Publications of the European University Institute

Department of Economics Working Paper Series

To Department of Economics **WP**
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
E-mail: publish@datacomm.iue.it
Italy

From Name . . . .

Address. . . . .

. . . .

. . . .

. . . .

. . . .

(Please print)

❒ Please enter/confirm my name on EUI Economics Dept. Mailing List
❒ Please send me a complete list of EUI Working Papers
❒ Please send me a complete list of EUI book publications
❒ Please send me the EUI brochure Academic Year 1995/96

Please send me the following EUI **ECO** Working Paper(s):

No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .

Date . . . . Signature . . . .

Working Papers of the Department of Economics Published since 1993

ECO No. 93/1
Carlo GRILLENZONI
Forecasting Unstable and Non-Stationary Time Series

ECO No. 93/2
Carlo GRILLENZONI
Multilinear Models for Nonlinear Time Series

ECO No. 93/3
Ronald M. HARSTAD/Louis PHLIPS
Futures Market Contracting When You Don't Know Who the Optimists Are

ECO No. 93/4
Alan KIRMAN/Louis PHLIPS
Empirical Studies of Product Markets

ECO No. 93/5
Grayham E. MIZON
Empirical Analysis of Time Series: Illustrations with Simulated Data

ECO No. 93/6
Tilman EHRBECK
Optimally Combining Individual Forecasts From Panel Data

ECO NO. 93/7
Víctor GÓMEZ/Agustín MARAVALL
Initializing the Kalman Filter with Incompletely Specified Initial Conditions

ECO No. 93/8
Frederic PALOMINO
Informed Speculation: Small Markets Against Large Markets

ECO NO. 93/9
Stephen MARTIN
Beyond Prices Versus Quantities

ECO No. 93/10
José María LABEAGA/Angel LÓPEZ
A Flexible Demand System and VAT Simulations from Spanish Microdata

ECO No. 93/11
Maozu LU/Grayham E. MIZON
The Encompassing Principle and Specification Tests

ECO No. 93/12
Louis PHLIPS/Peter MØLLGAARD
Oil Stocks as a Squeeze Preventing Mechanism: Is Self-Regulation Possible?

ECO No. 93/13
Pieter HASEKAMP
Disinflation Policy and Credibility: The Role of Conventions

ECO No. 93/14
Louis PHLIPS
Price Leadership and Conscious Parallelism: A Survey

ECO No. 93/15
Agustín MARAVALL
Short-Term Analysis of Macroeconomic Time Series

ECO No. 93/16
Philip Hans FRANSES/Niels HALDRUP
The Effects of Additive Outliers on Tests for Unit Roots and Cointegration

ECO No. 93/17
Fabio CANOVA/Jane MARRINAN
Predicting Excess Returns in Financial Markets

ECO No. 93/18
Iñigo HERGUERA
Exchange Rate Fluctuations, Market Structure and the Pass-through Relationship

ECO No. 93/19
Agustín MARAVALL
Use and Misuse of Unobserved Components in Economic Forecasting

ECO No. 93/20
Torben HOLVAD/Jens Leth HOUGAARD
Measuring Technical Input Efficiency for Similar Production Units: A Survey of the Non-Parametric Approach

*out of print

ECO No. 93/21
Stephen MARTIN/Louis PHLIPS
Product Differentiation, Market Structure and Exchange Rate Passthrough

ECO No 93/22
F. CANOVA/M. FINN/A. R. PAGAN
Evaluating a Real Business Cycle Model

ECO No 93/23
Fabio CANOVA
Statistical Inference in Calibrated Models

ECO No 93/24
Gilles TEYSSIÈRE
Matching Processes in the Labour Market in Marseilles. An Econometric Study

ECO No 93/25
Fabio CANOVA
Sources and Propagation of International Business Cycles: Common Shocks or Transmission?

ECO No. 93/26
Marco BECHT/Carlos RAMÍREZ
Financial Capitalism in Pre-World War I Germany: The Role of the Universal Banks in the Financing of German Mining Companies 1906-1912

ECO No. 93/27
Isabelle MARET
Two Parametric Models of Demand, Structure of Market Demand from Heterogeneity

ECO No. 93/28
Stephen MARTIN
Vertical Product Differentiation, Intra-industry Trade, and Infant Industry Protection

ECO No. 93/29
J. Humberto LOPEZ
Testing for Unit Roots with the k-th Autocorrelation Coefficient

ECO No. 93/30
Paola VALBONESI
Modelling Interactions Between State and Private Sector in a "Previously" Centrally Planned Economy

ECO No. 93/31
Enrique ALBEROLA ILA/J. Humberto LOPEZ/Vicente ORTS RIOS
An Application of the Kalman Filter to the Spanish Experience in a Target Zone (1989-92)

ECO No. 93/32
Fabio CANOVA/Morten O. RAVN
International Consumption Risk Sharing

ECO No. 93/33
Morten Overgaard RAVN
International Business Cycles: How much can Standard Theory Account for?

ECO No. 93/34
Agustín MARAVALL
Unobserved Components in Economic Time Series *

ECO No. 93/35
Sheila MARNIE/John MICKLEWRIGHT
"Poverty in Pre-Reform Uzbekistan: What do Official Data Really Reveal?"

ECO No. 93/36
Torben HOLVAD/Jens Leth HOUGAARD
Measuring Technical Input Efficiency for Similar Production Units: 80 Danish Hospitals

ECO No. 93/37
Grayham E. MIZON
A Simple Message for Autocorrelation Correctors: DON'T

ECO No. 93/38
Barbara BOEHNLEIN
The Impact of Product Differentiation on Collusive Equilibria and Multimarket Contact

ECO No. 93/39
H. Peter MØLLGAARD
Bargaining and Efficiency in a Speculative Forward Market

***

*out of print

ECO No. 94/1
Robert WALDMANN
Cooperatives With Privately Optimal Price Indexed Debt Increase Membership When Demand Increases

ECO No. 94/2
Tilman EHRBECK/Robert WALDMANN
Can Forecasters' Motives Explain Rejection of the Rational Expectations Hypothesis?

ECO No. 94/3
Alessandra PELLONI
Public Policy in a Two Sector Model of Endogenous Growth *

ECO No. 94/4
David F. HENDRY
On the Interactions of Unit Roots and Exogeneity

ECO No. 94/5
Bernadette GOVAERTS/David F. HENDRY/Jean-François RICHARD
Encompassing in Stationary Linear Dynamic Models

ECO No. 94/6
Luigi ERMINI/Dongkoo CHANG
Testing the Joint Hypothesis of Rationality and Neutrality under Seasonal Cointegration: The Case of Korea

ECO No. 94/7
Gabriele FIORENTINI/Agustín MARAVALL
Unobserved Components in ARCH Models: An Application to Seasonal Adjustment *

ECO No. 94/8
Niels HALDRUP/Mark SALMON
Polynomially Cointegrated Systems and their Representations: A Synthesis

ECO No. 94/9
Mariusz TAMBORSKI
Currency Option Pricing with Stochastic Interest Rates and Transaction Costs: A Theoretical Model

ECO No. 94/10
Mariusz TAMBORSKI
Are Standard Deviations Implied in Currency Option Prices Good Predictors of Future Exchange Rate Volatility?

ECO No. 94/11
John MICKLEWRIGHT/Gyula NAGY
How Does the Hungarian Unemployment Insurance System Really Work?

ECO No. 94/12
Frank CRITCHLEY/Paul MARRIOTT/Mark SALMON
An Elementary Account of Amari's Expected Geometry

ECO No. 94/13
Domenico Junior MARCHETTI
Procyclical Productivity, Externalities and Labor Hoarding: A Reexamination of Evidence from U.S. Manufacturing

ECO No. 94/14
Giovanni NERO
A Structural Model of Intra-European Airline Competition

ECO No. 94/15
Stephen MARTIN
Oligopoly Limit Pricing: Strategic Substitutes, Strategic Complements

ECO No. 94/16
Ed HOPKINS
Learning and Evolution in a Heterogeneous Population

ECO No. 94/17
Berthold HERRENDORF
Seigniorage, Optimal Taxation, and Time Consistency: A Review

ECO No. 94/18
Frederic PALOMINO
Noise Trading in Small Markets *

ECO No. 94/19
Alexander SCHRADER
Vertical Foreclosure, Tax Spinning and Oil Taxation in Oligopoly

ECO No. 94/20
Andrzej BANIAK/Louis PHLIPS
La Pléiade and Exchange Rate Pass-Through

ECO No. 94/21
Mark SALMON
Bounded Rationality and Learning; Procedural Learning

*out of print

ECO No. 94/22
Isabelle MARET
Heterogeneity and Dynamics of Temporary Equilibria: Short-Run Versus Long-Run Stability

ECO No. 94/23
Nikolaos GEORGANTZIS
Short-Run and Long-Run Cournot Equilibria in Multiproduct Industries

ECO No. 94/24
Alexander SCHRADER
Vertical Mergers and Market Foreclosure: Comment

ECO No. 94/25
Jeroen HINLOOPEN
Subsidising Cooperative and Non-Cooperative R&D in Duopoly with Spillovers

ECO No. 94/26
Debora DI GIOACCHINO
The Evolution of Cooperation: Robustness to Mistakes and Mutation

ECO No. 94/27
Kristina KOSTIAL
The Role of the Signal-Noise Ratio in Cointegrated Systems

ECO No. 94/28
Agustín MARAVALL/Víctor GÓMEZ
Program SEATS "Signal Extraction in ARIMA Time Series" - Instructions for the User

ECO No. 94/29
Luigi ERMINI
A Discrete-Time Consumption-CAP Model under Durability of Goods, Habit Formation and Temporal Aggregation

ECO No. 94/30
Debora DI GIOACCHINO
Learning to Drink Beer by Mistake

ECO No. 94/31
Víctor GÓMEZ/Agustín MARAVALL
Program TRAMO "Time Series Regression with ARIMA Noise, Missing Observations, and Outliers" - Instructions for the User

ECO No. 94/32
Ákos VALENTINYI
How Financial Development and Inflation may Affect Growth

ECO No. 94/33
Stephen MARTIN
European Community Food Processing Industries

ECO No. 94/34
Agustín MARAVALL/Christophe PLANAS
Estimation Error and the Specification of Unobserved Component Models

ECO No. 94/35
Robbin HERRING
The "Divergent Beliefs" Hypothesis and the "Contract Zone" in Final Offer Arbitration

ECO No. 94/36
Robbin HERRING
Hiring Quality Labour

ECO No. 94/37
Angel J. UBIDE
Is there Consumption Risk Sharing in the EEC?

ECO No. 94/38
Berthold HERRENDORF
Credible Purchases of Credibility Through Exchange Rate Pegging: An Optimal Taxation Framework

ECO No. 94/39
Enrique ALBEROLA ILA
How Long Can a Honeymoon Last? Institutional and Fundamental Beliefs in the Collapse of a Target Zone

ECO No. 94/40
Robert WALDMANN
Inequality, Economic Growth and the Debt Crisis

ECO No. 94/41
John MICKLEWRIGHT/Gyula NAGY
Flows to and from Insured Unemployment in Hungary

*out of print

ECO No. 94/42
Barbara BOEHNLEIN
The Soda-ash Market in Europe: Collusive and Competitive Equilibria With and Without Foreign Entry

ECO No. 94/43
Hans-Theo NORMANN
Stackelberg Warfare as an Equilibrium Choice in a Game with Reputation Effects

ECO No. 94/44
Giorgio CALZOLARI/Gabriele FIORENTINI
Conditional Heteroskedasticity in Nonlinear Simultaneous Equations

ECO No. 94/45
Frank CRITCHLEY/Paul MARRIOTT/ Mark SALMON
On the Differential Geometry of the Wald Test with Nonlinear Restrictions

ECO No. 94/46
Renzo G. AVESANI/Giampiero M. GALLO/Mark SALMON
On the Evolution of Credibility and Flexible Exchange Rate Target Zones

***

ECO No. 95/1
Paul PEZANIS-CHRISTOU
Experimental Results in Asymmetric Auctions - The 'Low-Ball' Effect

ECO No. 95/2
Jeroen HINLOOPEN/Rien WAGENVOORT
Robust Estimation: An Example

ECO No. 95/3
Giampiero M. GALLO/Barbara PACINI
Risk-related Asymmetries in Foreign Exchange Markets

ECO No. 95/4
Santanu ROY/Rien WAGENVOORT
Risk Preference and Indirect Utility in Portfolio Choice Problems

ECO No. 95/5
Giovanni NERO
Third Package and Noncooperative Collusion in the European Airline Industry

ECO No. 95/6
Renzo G. AVESANI/Giampiero M. GALLO/Mark SALMON
On the Nature of Commitment in Flexible Target Zones and the Measurement of Credibility: The 1993 ERM Crisis

ECO No. 95/7
John MICKLEWRIGHT/Gyula NAGY
Unemployment Insurance and Incentives in Hungary

ECO No. 95/8
Kristina KOSTIAL
The Fully Modified OLS Estimator as a System Estimator: A Monte-Carlo Analysis

*out of print